Acoustic Force Spectroscopy (AFS)

Kamsma, D.

2018

document version
Publisher's PDF, also known as Version of record

Chapter 6 |

Unravelling the folding energy landscape of short-handle DNA hairpins using Acoustic Force Spectroscopy

Abstract

The way biomolecules fold is crucial for their function. The experimental tool to investigate folding mechanics is force spectroscopy, which allows exploration of the free energy landscape of folding under the action of mechanical force applied to the biomolecule of interest. To map the whole energy landscape, many individual measurements have to be performed, over a wide range of forces. Here we show that Acoustic Force Spectroscopy (AFS) is ideally suited for this kind of measurements by measuring the folding free energy landscape of a simple, well-characterized DNA hairpin. We demonstrate that folding and unfolding of many hairpins can be measured in parallel, while applying force loading rates over 5 orders of magnitude. We use the standard Bell-Evans theory, as well as the Continuous Effective Barrier Approach (CEBA) to obtain values for the free energy difference between the states, the height and position of the transition state.

6.1 Introduction

Biomolecules like DNA, RNA and proteins execute specific tasks that are dictated by the sequence of their polymer chains and by their conformation. A DNA molecule, for example, can only be read or duplicated when its strands are melted and the sequence is available for transcription[94]. An RNA molecule manufactured by this reading process can exist in activated and non-activated states[95], with different conformation. Only activated messenger RNA molecules are translated, resulting in proteins with specific function related to their conformation[96–98]. Understanding the folding processes of polymer chains has been one of the main challenges of molecular biophysicists for many years and is known as the "folding problem"[99].

The folding problem is particularly large for proteins. A polymer chain of 150 amino acids will have on the order of ~10^{80} conformational states[100]. The problem is for many proteins even more complex, since they do not always fold to their lowest energy conformation, but in many cases certain higher energy conformations are stabilized by environmental factors, such as the ribosomes[101] or chaperones[102,103]. Protein folding and missfolding are of particular interest not only from a physics perspective, but also from a medical perspective, since misfolded proteins are at the heart of diseases, such as Parkinson's, Alzheimer's[104] and many others[105].

One approach to increase our understanding of biomolecule folding are computer simulations. One notable example is the computer game "foldit"[106], which challenges players to find the optimal configuration of a protein. Biomolecular folding has also been studied experimentally, for example using fluorescence spectroscopy[107,108] and stopped-flow mixing[109,110]. Single-molecule force spectroscopy experiments, making use of Atomic Force Microscopy (AFM)[111] or Optical Tweezers (OT)[112] have allowed direct observation of folding and unfolding events of individual biomolecules. Furthermore, these approaches allowed probing of the free energy landscapes constituted by folded, unfolded and transitions states and of the kinetics of transitions between the states. In many cases, the Bell-Evans theory[113] is used to find the free energy difference between the states and the location of the barrier (**Figure 6.1a**). To this end, many transitions between the states need to be probed at different force loading rates. AFM and OT have, in this respect, as main drawback that only one molecule can be probed at the time, limiting data throughput. AFS, a novel force-spectroscopy approach using acoustic standing waves as force generator (**Chapter 2**), on

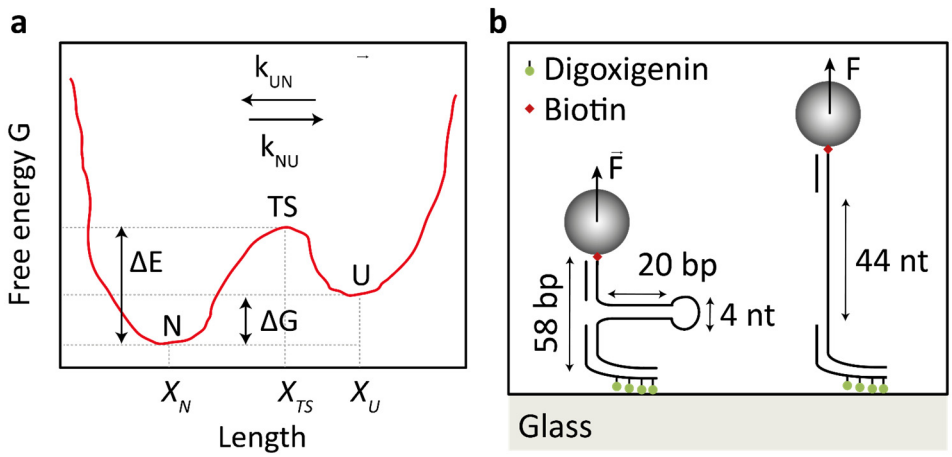


Figure 6.1 | Sketch of the folding free-energy landscape of the CD4 hairpin and the AFS assay
(**a**) One-dimensional representation of the free-energy landscape of CD4 modeled as a two-state system[114]. Native (N), transition (TS) and unfolded (U) states are drawn with arbitrary energies. (**b**) Structure of the CD4 DNA hairpin, consisting of a 20 base-pair stem with a 4 base tetraloop on its end (see **Supplementary figure 6.1** for detailed description of the synthesis).

the other hand, allows simultaneous measurements on more than hundreds of biomolecules in parallel and thus promises a much higher data throughput. High data throughput can be achieved with magnetic tweezer[115] as well, however AFS also allows a very wide range of force loading rates, making it, in principle, an ideal method to unravel the folding free energy landscape of biomolecules.

As a proof of concept for the application of AFS to the folding of biomolecules, we determined the free energy landscape of a widely studied DNA hairpin, CD4 (**Figure 6.1b**). This hairpin, one single-stranded DNA strand, consists of two sequence regions (of 20 bases) that are complementary to each other. In the hairpin's native state, these regions interact with each other by base pairing, resulting in a closed state. The interaction between these complementary regions can be overcome by the application of force to the ends of the DNA strand (**Figure 6.1b**). The short handles of the CD4 hairpin[116] should result in high spatial and temporal resolution. In a free-energy diagram, the folding and unfolding of the hairpin can be described as a two-state system (**Figure 6.1a).** We show that the folding and unfolding of this hairpin under external load can be measured with AFS, with the key advantage that multiple molecules can be monitored at the same time. We use the Bell-Evans theory and CEBA to

analyse our data and successfully map the free-energy landscape of CD4 hairpin.

6.2 Methods and materials

In this chapter, measurements are performed with the AFS system described in (**Chapter 3**). We use a short-handle DNA hairpin construct, fabricated in the laboratory of Felix Ritort (Universitat de Barcelona). The surface-tethering protocol used is a modification of the one described in (**Chapter 4**), required to prevent sticking of the microspheres to the glass surface. In addition, a commercial temperature control unit was employed.

6.2.1 CD4 hairpin with short handles

Three different synthetic oligonucleotides are used to make the short handle DNA hairpins (**Supplementary figure 6.1a**). The splint (the complimentary parts of the handle) works as a scaffold for DNA handle A and handle B. Both handles include one handle and part of the hairpin. Handle A has a biotin at the end, while handle B first had to be digoxigeninylayeted using an oligonucleotide tailing kit (Roche; 03353583910) (**Supplementary figure 6.1b**). Milli-Q water (8 µL), oligonucleotide (1 µL, 100 µM), Reaction Buffer x5 (4 µL), $CoCl_2$ (4 µL), dATP (1 µL), Dig-dUTPare (1 µL) and terminal transferase (1 µL, enzyme) are mixed in an Eppendorf tube in subsequent order and incubate for 15 minutes at 37˚C. After, 2 µL EDTA (0.2 M, pH8.0) was added. A Qiaquick Nucleotide Purification Kit (Quiagen) was used and 50 µL of the modified oligonucleotide (2 pmol/µL) was obtained. Next, both handles were annealed separately to the splint (**Supplementary figure 6.1c**) and after to each other (**Supplementary figure 6.1c**). Both handles (5 pM, final concentration) were mixed with the SPLINT (5 pM final concentration), 1 M Tris (1 µL, pH 7.5) and 5 M NaCl (0.5 µL) and Milli-Q water (added to a final volume of 15 µL) and incubated for 4 hours at 42˚C. Subsequently, the two annealing products were mixed together and incubated for 30 minutes at 23˚C. The buffer was exchanged by dialysis for 30 minutes (50 mL 5 mM NaCl, 20 mM Tris pH7.5). In the last step, the backbone of the DNA was ligated together (**Supplementary figure 6.1d**). The DNA annealing product was mixed with ATP (2 µL, 10 mM), Ligation Buffer x10 (2 µL), T4 DNA ligase (2 µL) and Milli-Q water (supplemented till a final volume of 20 µL) and incubated for 16 hours (typically overnight) at 16 ˚C. The buffer was exchanged by dyalisis (50 mL of 10 mM NaCl, 10 mM Tris pH7.5, 1 mM EDTA; 30 minutes). In the final last step, the sample was heated to 70˚C for 10 minutes, to thermally inactivate the ligation enzyme. We measured a concentration of 1.6 mg/mL. The final product was aliquoted

and kept at -20˚C for long time storage.

6.2.2 Surface tethering protocol

To reduce non-specific interactions with the surface, the tethering protocol used in **Chapter 4** was slightly modified. In this chapter only the bottom force flow cells were used, described in **Chapter 3**. For thorough cleaning, the flow cell was incubated with NaOH for 30 minutes and afterwards flushed with air and Milli-Q water repeatedly. Next, the flow cell was made hydrophobic using a siliconizing reagent (Sigma Aldrich, SL2 SIGMA), incubated for 1 minute. Subsequently, the residues of the silicon were removed with acetone and the flow cell was thoroughly washed with Milli-Q water and air. The flow cell was incubated with Anti-digoxigenin (Roche, Cat. No. 11 333 089 001) for 20 minutes, followed by 30 minutes of incubation with PBS (138 mM NaCl, 2.7 mM KCl and 10 mM phosphate (pH 7.4)), supplemented with EDTA (5 mM), BSA (0.2% w/v), Pluronics F127 (0.2% w/v) and Sodiumazide (0.6 mg/mL), which we will refer to as measuring buffer. Surfaces were used for multiple subsequent measurements.

Microspheres were pre-incubated with the DNA hairpin, before flushing them into the flow cell. 20 μL of 4.5 μm (diameter) streptavidin-coated polystyrene microspheres (0.15% w/v; SVP-40-5, Spherotech) were mixed with 2 μL of digoxigenin labeled microspheres (3.05 μm; **section 4.3.3**), 20 μL of measuring buffer and 1 μL of DNA hairpin (1.6 μg/mL), and tumbled for 20 minutes. The microsphere sample was cleaned by adding 1 mL of measuring buffer, followed by spinning down (2000 × g for 2 minutes) and discarding the supernatant. This was repeated three times and the final product was dissolved in 50 μL measuring buffer.

10 μL of the DNA-hairpin and digoxigenin labeled microspheres were flushed in the flow cell and left to incubate for 10 minutes. Subsequently, unbound microspheres were gently flushed out the flow cell with measuring buffer (**Chapter 4, note 4**) and measurements were started. Afterwards, all microspheres were removed with strong flow and new microspheres were flushed in to start a new measurement. Typically, a single surface-treated flow cell was used for a couple of weeks.

6.2.3 Temperature control

In this chapter, a temperature control system (LUMICKS; AFS-TC) was used to increase the temperature up to 45˚C. The temperature control system is designed as an inset in the standard chip holder of the AFS. Two heating elements heat up the whole flow cell, while a temperature sensor measures

the temperature in the middle of the flow cell. A proportional–integral–derivative (PID) controller is applied to keep the temperature at a set point. This system can only heat and not cool the sample.

6.3 Theoretical background of a two-state model

In **Figure 6.1a** an example free-energy landscape of a two-state system is shown. The free energy of the system changes according to the reaction coordinate, in our case the end-to-end length of the molecule. From thermodynamics we know that a system will always try to reach to the lowest energy state, helped by thermal energy. The ratio of the probabilities of finding the system in a given state (say the folded one) with respect to another one (e.g. the unfolded one) depends on the free-energy difference of the states (ΔG). To switch between the two states, a barrier (B) has to be crossed i.e. the transition state (TS) which is defined as the highest energetic state separating the two states along the path of reaction coordinate(s). The free energy landscape of molecular transitions and thus the TS typically depends on temperature[117], salt concentration[118] and force applied[119]. While the height of TS is directly related to the rate of transitioning from one state to the next. With force spectroscopy, rates can be determined relatively straightforward. In the Arrhenius equation, the kinetic rate (k) for a first-order chemical reaction is described by a single exponential decay function:

$$k = k_0 \exp\left(-\frac{\Delta E}{k_B T}\right) \tag{6.1}$$

With k_0 is the reaction rate constant, also called the pre-exponential factor or the attempt frequency, ΔE is the activation energy, the energy difference between the starting state and the TS, k_B is the Boltzmann constant and T the temperature. In case the ΔE of the transition is very high (compared to $k_B T$), the molecule will be trapped in its starting state. To probe the energy landscape for these kind of cases, a mechanical load can be applied to the molecule to push it out of its stable starting state. The effect of the applied load to the energy landscape can be considered to resolve the energy landscape of the molecule. To this end, we will apply the widely used Bell-Evans theory[113,120] and the more recently developed CEBA theory[119,121] to further interpret ΔE, TS and ΔG in terms of a two-state system that is perturbed by applying external force.

6.3.1 Bell-Evans theory

According to the Bell-Evans theory[113,120], the transition state in a two-state

system is characterized by two parameter: the barrier height B_{BE} and the barrier position x_{BE}. Applying an external force to the two-state system causes the energy landscape to tilt, resulting in an effective activation energy that is the sum of the barrier height B_{BE} and a force-dependent term $-f\ x_{BE}$ (**Supplementary figure 6.2**):

$$\Delta E(f) = B_{BE} - f \cdot x_{BE} \tag{6.2}$$

By filling this expression for ΔE in **equation 6.1** we obtain for the unfolding and folding rates (k_{NU} and k_{UN}):

$$k_{NU}(f) = k_m * \exp\left(\frac{f \cdot x_{N-TS}}{k_B T}\right) \tag{6.3}$$

$$k_{UN}(f) = k_m * exp\left(\frac{\Delta G - f \cdot x_{TS-U}}{k_B T}\right) \tag{6.4}$$

$$k_m = k_0 * \exp\left(-\frac{B_{BE}}{k_B T}\right) \tag{6.5}$$

Here k_m is the unfolding rate at zero force, $x_{N\text{-}TS}$ and $x_{TS\text{-}U}$ are the distances between the native and the transition state, and the transition and the unfolded state, respectively. Measuring the rates over a range of different forces allows calculating k_m, x_N and x_0. In practice, the natural logarithms of the rates are fitted with a straight line:

$$\ln\left(k_{NU}(f)\right) = a_N + f \cdot b_N \tag{6.6}$$

$$\ln\left(k_{UN}(f)\right) = a_U + f \cdot b_U \tag{6.7}$$

Where a_N and a_U are the intercepts of the fit and b_N and b_U the slopes of the lines fit to the data of the native and unfolding state, respectively. From these fits, values for k_m, ΔG, $x_{N\text{-}TS}$ and $x_{TS\text{-}U}$ can be retrieved as follows:

$$k_m = \exp(a_N) \tag{6.8}$$

$$\Delta G = (a_U - a_N) \cdot k_B T \tag{6.9}$$

$$x_{N-TS} = b_N \cdot k_B T \tag{6.10}$$

$$x_{TS-U} = b_U \cdot k_B T \tag{6.11}$$

Note that the sum of $x_{N\text{-}TS}$ and $x_{TS\text{-}U}$ is equal to the distance between the folded and unfolded states ($x_{N\text{-}U}$). The force at which folding and unfolding rates are

equal is the coexistence force (f_c) :

$$f_c = \frac{\Delta G}{x_{N-TS} + x_{TS-U}} \tag{6.12}$$

Using Bell-Evans theory, the distance to the transition state can be determined as well as ΔG and k_m. Note that, k_m depends on both B_{BE} and k_0, therefore B_{BE} and k_0 cannot be obtained in a straightforward way by applying Bell-Evans theory.

6.3.2 Continuous Effective Barrier Approach (CEBA)

The Bell-Evans theory predicts that, for a two-state system, the natural logarithm of the rates is linearly dependent on the applied force (**equation 6.6** and **6.7**). In experiments, deviations from linear behavior have been observed at forces substantially larger or smaller than the coexistence force[119]. This effect cannot be explained the by the Bell-Evans or other theories like the one from Dudko-Hummer-Szabo[122]. CEBA, in contrast, assumes a different, non-linear force-dependent barrier $B(f)$:

$$\Delta E = B(f) \tag{6.13}$$

In the case of a DNA hairpin, the $B(f)$ can be theoretically predicted using a discretized version of the Kramers' 1D diffusive model[119,123]:

$$B(f) = k_b T \ln\left[\sum_{n=0}^{N}\sum_{n'=0}^{n} \exp\left(\frac{\Delta G_n(f) - \Delta G_{n'}(f)}{k_b T}\right)\right] \tag{6.14}$$

$B(f)$ is a double sum of all the possible configurations (n) of the hairpin and N is the total number of base pairs in the stem. For the hairpin used in this study, $\Delta G(f)$ can be described by[119]:

$$\Delta G(f) = \Delta G^0 + \Delta G^{ssDNA}(f) + \Delta G^d(f) \tag{6.15}$$

ΔG^0 is the free energy of the formation at zero force, $\Delta G^{ssDNA}(f)$ is an elastic contribution accounting for the reversible mechanical work required to stretch the ssDNA after unzipping, and $\Delta G^d(f)$ an elastic term to account for the hairpin's orientation along the axis of force (see below).

ΔG^0 is calculated by summing all free energies of the individual base pairs using the Nearest-Neighbor (NN) model[124] (which only takes into account the base pair itself and its nearest neighbor (in the 3'-5' direction)) and adding a contribution of the loop. Values used in the NN model are taken from single-

molecule DNA unzipping experiments[125], for the complementary stem and for the loop, respectively.

$\Delta G^{ssDNA}(f)$ is the reversible mechanical work required to stretch the ssDNA from a random configuration to a stretched equilibrium end-to-end distance (x_L), at force f. x_L is calculated using the Worm-Like Chain (WLC) model[126]:

$$f = \frac{k_b T}{P}\left(\frac{1}{4(1 - x_L/L_c)^2} - \frac{1}{4} + \frac{x_L}{L_c}\right) \tag{6.16}$$

with P the persistence length and L_c the contour length of the ssDNA. To obtain the $\Delta G^{ssDNA}(f)$, the WLC function is integrated from 0 to the force applied f. Parameters used are 0.59 nm/base for the interphosphate distance and 1.35 nm for the persistence length[127].

$\Delta G^{d}(f)$ is the energy contribution of the hairpin's orientation along the axis of force. In other words, when the hairpin is in the closed state, there is a small contribution to the distance between the two handles that is related to diameter of the double stranded DNA (sdDNA) helix (d). $\Delta G^{d}(f)$ can be expressed by[116]:

$$\Delta G^{d}(f) = k_b T \ln\left[\frac{k_B T}{f d(f)} \sinh\left(\frac{f d(f)}{k_B T}\right)\right](1 - \delta) \tag{6.17}$$

with δ equal to 1 for the closed state and 0 otherwise. $d(f)$ is close to the diameter of the dsDNA part of the hairpin ($d(f)$, here taken to be equal to 2 nm[119]) for the applied force range.

Now $\Delta G(f)$ can be calculated and plugged into **equation 6.14** to obtain expressions that relate the experimentally obtained folding and unfolding rates to the calculated $B(f)$ (using **equations 6.13** and **6.1**):

$$k_{NU}(f) = k_0 \exp\left(-\frac{B(f)}{k_B T}\right) \tag{6.18}$$

$$k_{UN}(f) = k_o \exp\left(-\frac{B(f) - \Delta G(f)}{k_B T}\right) \tag{6.19}$$

$$\frac{B(f)}{k_B T} = \ln[k_0] - \ln[k_{NU}(f)] \tag{6.20}$$

$$\frac{B(f)}{k_B T} = \ln[k_0] - \ln[k_{UN}(f)] + \frac{\Delta G(f)}{k_B T} \tag{6.21}$$

$$= \ln[k_0] - \ln[k_{UN}(f)] + \frac{\Delta G^0 + \Delta G^{ssDNA}(f) + \Delta G^d(f)}{k_B T}$$

These equations show that k_0 can be calculated from the offset between theoretically calculated $B(f)$ and the measured unfolding rates (eq. 6.20). In addition, ΔG^0 can be obtained in a similar way from $B(f)$ and the measured folding rates, taking into account the elastic contributions (eq. 6.21). This determination of ΔG^0 can serve as a verification of the theoretically calculated $B(f)$.

6.4 Results

To study the force-dependence of the folding and unfolding of the DNA-hairpin CD4 using AFS, we tethered hairpins between the surface of the AFS chip and microspheres. A typical field of view contains about 40 hairpin-tethered microspheres and 5 microspheres that are directly attached to the surface (**Supplementary figure 6.4**). The xy and z coordinates of the microspheres are tracked in time, as explained before (**section 4.3.5**). The positions of at least two surface-attached microspheres are used to correct for drift in the instrument. To obtain the end-to-end length of the CD4 DNA constructs, Pythagoras' theorem is used (**section 4.3.6**). In this way, hopping between the native or unfolded state of the DNA hairpin could be observed (**Figure 6.2a)**. Both the power spectrum method (**section 2.4.4**) and the equipartition theorem[31] are used to calibrate forces acting on individual tethers. Tethers were only used for further analysis when both methods led to the same result.

6.4.1 Constant force spectroscopy

The most straightforward approach to acquire folding and unfolding rates for the hairpins switching between native (folded) and unfolded states is to apply a constant force with AFS. Since the force acting on the tethered microspheres in an AFS chip depends only minimally on the height of the sphere in the chip, AFS effectively works as a force clamp. An example of an experiment where the force is clamped is shown in **Figure 6.2a.** The hairpin can be observed to hop between the native and the unfolded states, resulting in tether-length changes of ~20 nm, corresponding to the tether lengthening with 44 single stranded nucleotides. Single tethers are selected based on the total end-to-end length and the distance between native and unfolded state.

To determine the end-to-end lengths of the tether in both states more accurately, all end-to-end lengths between 10 and 50 minutes of the data in

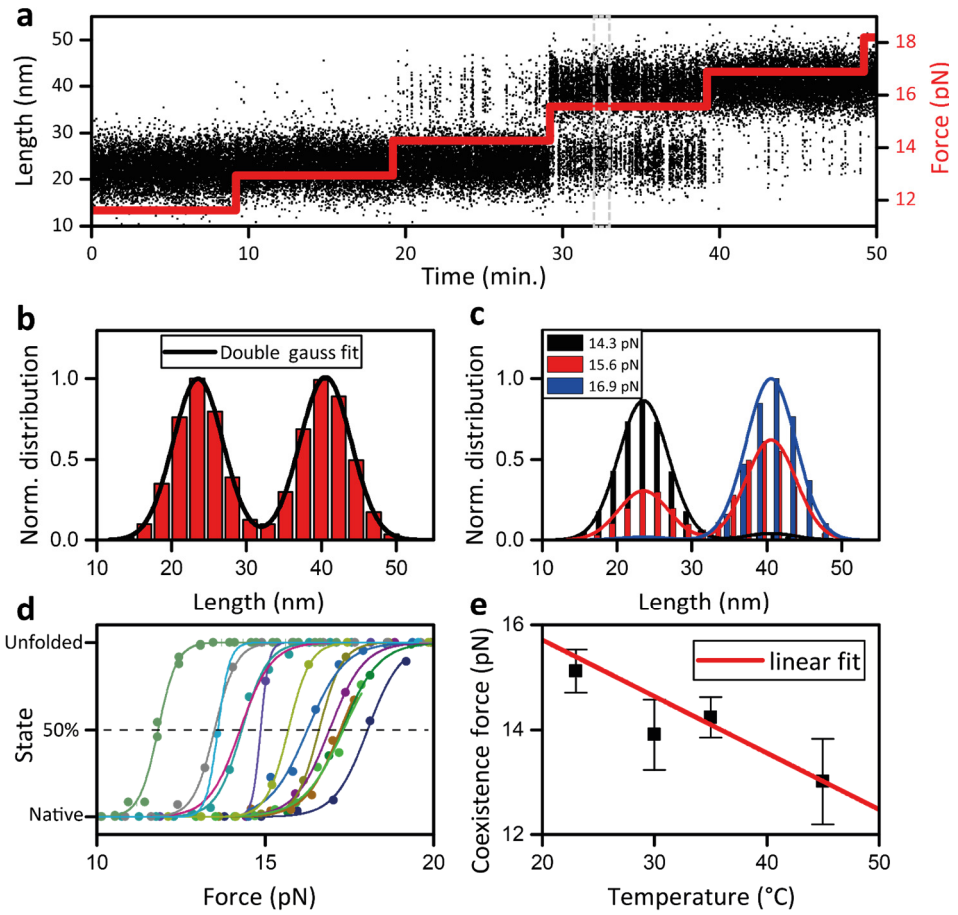


Figure 6.2 | Quantification the CD4 DNA hairpin state under constant tension
(**a**) Time trace of the measured end-to-end lengths of the CD4 construct at different tensions close to the coexistence force. Data is recorded with 60 Hz; for clarity, 2.5% of the complete trace is shown. (**b**) Normalized histogram of the end-to-end lengths (all data points in **a** between 10 to 50 minutes are included), fitted with a double Gaussian function. Fit values of the peaks are 23.51658 ± 0.0005 and 40.56055 ± 0.00005 and for the width 6.67477 ± 0.00009 and 6.83646 ± 0.00009 (values ± s.e.m., obtained from the fit). (**c**) Normalized histograms of the end-to-end lengths for shorter intervals in **a** where the force was kept constant (as indicated). Data is fitted with the same double Gaussian function as **b**, using only the ratio between the peaks as a fit value. Corresponding fit values are 4.5 ± 0.5, 67 ± 2 and 98.1 ± 1.3 % for 14.3, 15.6 and 16.9 pN, respectively (values ± s.e.m., obtained from the fit). (**d**) Fraction of time spent in the unfolded state as a function of applied force, measured for 13 individual hairpins. Data is fitted with a sigmoidal function and the average coexistence forces is 15.1 ± 0.4 pN (average ± s.e.m., obtained from the fit). (**e**) Coexistence force obtained at different temperatures fitted with a straight line. Fit values yield for the intercept 17.9 ± 0.6 pN and for the slope -0.11 ± 0.02 pN/°C (values ± s.e.m.). For each temperature, the acoustic resonance frequency was determined and adjusted (**Supplementary figure 6.3**)

Figure 6.2a are plotted in a histogram and fitted with a double Gaussian function (**Figure 6.2b**). The standard deviation of both Gaussians is ~3.4 nm, which is directly related to the accuracy of determining the end-to-end length of CD4. Next, segments of the data corresponding to a particular set force are fitted with a double Gaussian function, with fixed values of the Gaussian peak positions and the widths (as obtained from the larger data set). At different forces, different ratios of the native and unfolded states can be observed (**Figure 6.2c** and **d**). The force-dependence of this ratio has a sigmoidal form, with 50% of its points corresponding to the coexistence force of the DNA hairpin (**Figure 6.2d**). The coexistence force is temperature dependent, decreasing linearly with temperature, as expected from **equation 6.1**. Unexpectedly, the coexistence force varies substantially between hairpins[116]. We attribute this variance to imprecisions in our force calibration, which could be caused by variations in microsphere size and shape. Pending improvements to our force calibration method, we set the coexistence force of each hairpin to 15 pN (at 23°C) as a means of force calibration. Around the coexistence force, the hairpin frequently hops from one state to the other. To obtain insight in the kinetics of hairpin folding and unfolding, we determined dwell-time distributions (**Figure 6.3a**). Dwell times were obtained by thresholding the length-time trace, using the average of the average end-to-end lengths of the two states (at 32 nm, **Figure 6.2b**). Dwell times are exponentially distributed, as expected for a simple two-state system[128,129] (**Figure 6.3.b,c**). The distributions (**Figure 6.3b** and **c**) are fitted with a mono exponential decay function, yielding the folding and unfolding rates of the states. The rates depend strongly on the force applied (**Figure 6.3d**). Using the Bell-Evans theory, $x_{N\text{-}TS}$, $x_{TS\text{-}U}$, ΔG and F_c can be determined (**section 6.3.1**).

6.4.2 Dynamic force spectroscopy

In a force-clamp experiment, the system studied is kept in equilibrium. Uncovering the broader free energy landscape requires pushing the system out-of-equilibrium, which can be done by changing the force applied to the system with time: dynamic force spectroscopy. At high loading rates, the system can be pushed out-of-equilibrium, which can result in additional unfolding or refolding pathways (as is done in **Figure 2.4**). For the simple two-state system used here, this is, however, not expected. For this hairpin, using these high loading rates, allows probing the system over a broader range of forces than accessible in constant force experiments. This allows us to verify whether the CD4 hairpin behaves as an ideal two-state system also when pushed further out of equilibrium.

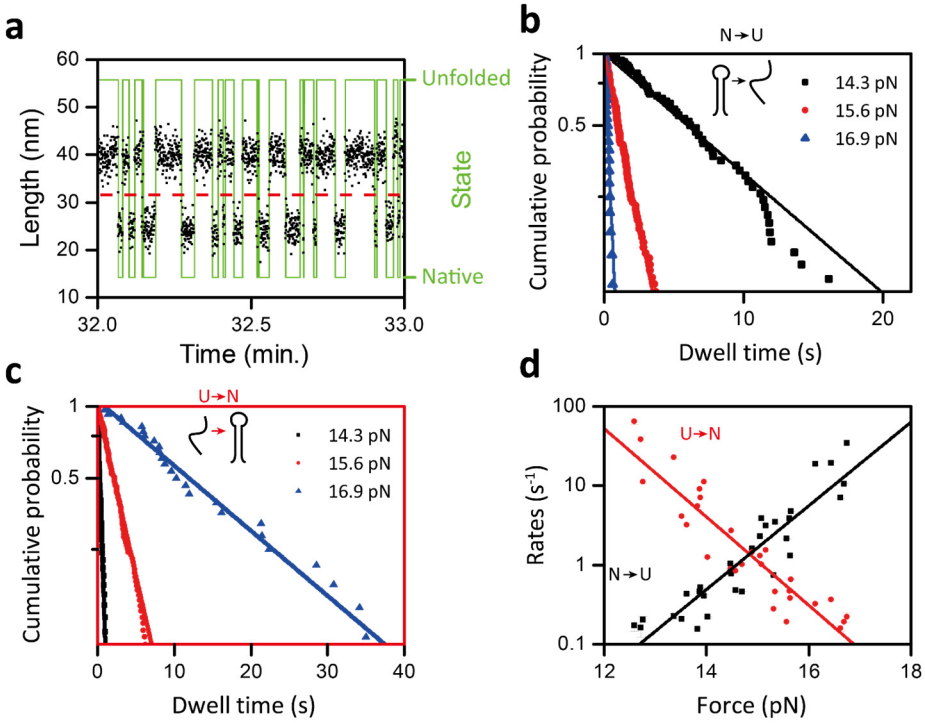


Figure 6.3 | Resolving the kinetics of the native and the unfolded state

(**a**) Zoom in of the end-to-end length time trace (**Figure 6.2a**, grey dashed line); binning of 2 time points is applied. Green line shows the state of the molecule determined by thresholding, using a value (red dashed line, at 32 nm) obtained from the average of the two gausssian peaks of **Figure 6.2b**. (**b**-**c**) Cumulative probalility distribution of the dwell times of the native (**b**) and the unfoled state (**c**), fitted with mono exponential decay functions (straigth lines). Fit values of the kinetic rates are 15.8 ± 0.4, 3.06 ± 0.02 and 0.445 ± 0.010 seconds for the native state (**b**) and 0.328 ± 0.012, 1.494 ± 0.008 and 8.62 ± 0.15 seconds for the unfolded state (**c**), while appling 14.3, 15.6 and 16.9 pN, respectively (values ± s.e.m., obtained from the fit). (**d**) Folding (red) and unfolding rates (black) as a function of the applied force, measured for 10 individual hairpins. The natural logarithm of the rates are fitted with straight lines, yielding 19.3 ± 1.5 and -17.7 ± 1.4 s^{-1} for the intercepts and 1.21 ± 0.10 and -1.28 ± 0.10 $s^{-1}pN^{-1}$ for the slopes of unfolding and folding, respectively (values ± s.e.m., obtained from the fit).

To apply dynamic force spectroscopy using AFS, the force applied is ramped up linearly by applying a square-root voltage ramp to the piezo on the AFS chip (**Supplementary figure 2.6**). Note that, in contrast to optical tweezers, where the typical length of the tethered molecule is changed linearly while the resulting force is measured, in AFS the force applied to the tether is changed linearly while the tether length is measured. Force-extension curves showing unfolding and folding of the DNA hairpin upon ramping up and down load, respectively, are shown in **Figure 6.4.** To extract values for ΔG,

x_{U-TS} and x_{TS-F} from this dynamic force data, the force of the first signature of unfolding (f_u) in the forward force stretching curve and the force of the first signature of folding (f_f) in the backward stretching curve have to be determined. f_u and f_f are determined by taking the derivative of the displacement and extracting the first peak correlated to the first folding or unfolding event (an example of one stretching curve is shown **Supplementary figure 6.5**). As expected, we observe that the spread in f_u and f_f is increasing with the increase in the force loading rate (**Figure 6.4b**). Next, we extract folding and unfolding rates from the experimentally measured distributions of f_u and f_f (**Figure 6.4b**). To this end, we analyze the distributions by converting them to cumulative probability distributions. Rates are extracted from these distributions assuming a first-order Markov process[123,130], as follows:

$$k_{N\to U}(f) = -r\frac{1}{P_{NU}}\frac{dP_{NU}(f)}{df} \tag{6.22}$$

$$k_{U\to N}(f) = -r\frac{1}{P_{UN}}\frac{dP_{UN}(f)}{df} \tag{6.23}$$

Where $P_{NU}(f)$ and $P_{UN}(f)$ are the cumulative survival probabilities of the states and r is the force loading rate. Using the formulas above, we observe that the rate constants as a function of force applied show the same straight line, independent of loading rate (**Figure 6.4c**). Using the Bell-Evans theory, values for ΔG, x_{N-TS} and x_{TS-U} can be calculated, which agree well with the values obtained from constant force measurements (**Figure 6.3c** and **Figure 6.4c**, **Table 6.1**).

	$x_{N-TS}(nm)$	$x_{TS-U}(nm)$	$\Delta G(k_bT)$	$f_c(pN)$
Constant force	5.0 ± 0.4	5.2 ± 0.4	37 ± 4	15 ± 4
Dynamic force	4.27 ± 0.14	4.41 ± 0.23	31 ± 2	15 ± 3

Table 6.1 | Fit results of the kinetic rates.

Using Bell-Evans theory for constant clamps and dynamic forces ramps. Values are calculated using **equation 6.8-12** and obtained fit values of **Figure 6.3d** and **Figure 6.4c** (± are s.e.m. obtained from the fit).

So far, we have not acquired information on the height of the barrier between native and unfolded conformation, since the Bell-Evans theory only

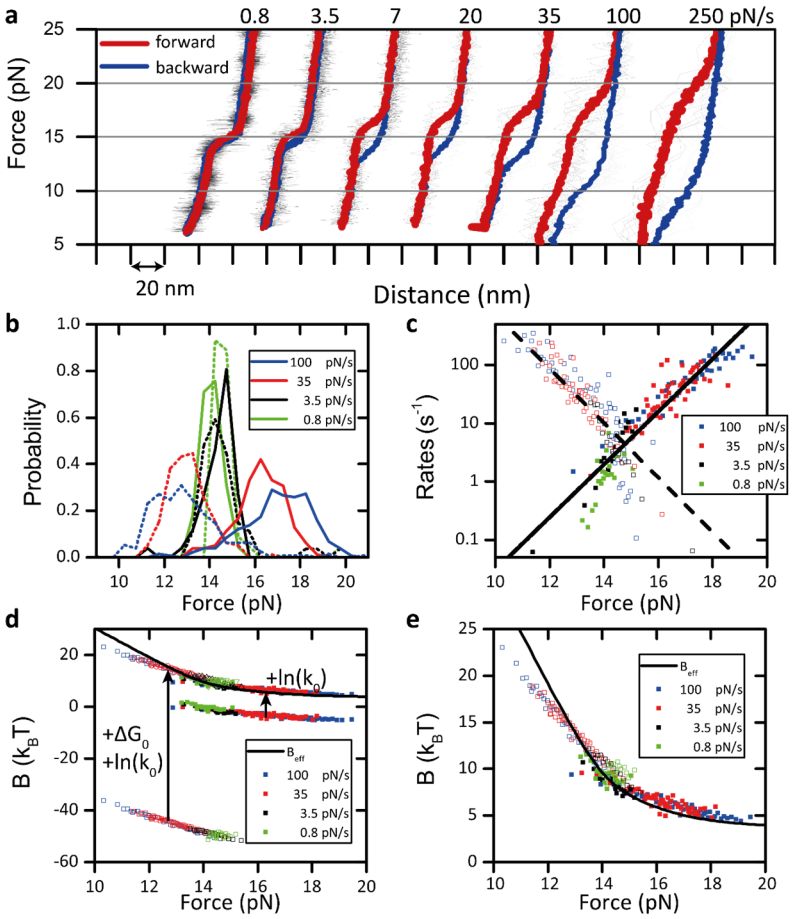


Figure 6.4 | Dynamic force spectroscopy experiments of CD4 hairpins
(a) Examples of individual unfolding and folding force-extension curves (thin grey lines) of the DNA hairpin for different loading rates, ranging from 0.8 to 250 pN. For each loading rate, curves (>100) are averaged and binning of 50 data points is applied, resulting in the forward (thick red) and backward (thick blue) curve showing increasing hysteresis with increasing loading rate. Data is recorded at 60 (for ≤ 35 pN/s) and 250 Hz (for ≤ 100 pN/s). Orange arrows indicate the part where the forward and the backward-stretching curves do not overlay due to the time response of the system. (**b**) Histograms of the forces of the first unfolding (solid line) and folding (dotted line) events in individual force-extension curves (see **Supplementary figure 6.5**). (**c**) Unfolding (solid squares) and folding rates (open squares) calculated from the data in **b**, using **equations 6.22** and **6.23**. The natural logarithms of the folding and unfolding rates are fitted with straight lines, yielding fit values of -13.9 ± 0.6 and 17.4 ± 0.8 s^{-1} for the intercept and 1.04 ± 0.03 and -1.08 ± 0.06 $s^{-1}pN^{-1}$ for the slope, for the native and the unfolded state, respectively (values ± s.e.m., obtained from the fit). (**d**) Theoretically predicted B(*f*) of the CD4 hairpin (black solid line; obtained from **equation 6.14**). Data is represented in four ways. Measured $-\ln(k_{UN})$ and measured $-\ln(k_{UN}) + \ln(k_0)$ are shown as solid squares. Measured $-\ln(k_{NU})$ (bottom of graph) and measured $-\ln(k_{NU}) + \ln(k_0) - \Delta G^d$ + calculated ΔG^{ssDNA} (top of graph) are shown as open squares. Data is shifted to overlay B(*f*) (as discussed in **section 6.3.2**), yielding a best fitting offset of 49.5 ± 1.5 and 9.7 ± 0.6 k_BT for ΔG^0 and $\ln(k_0)$, respectivly. (**e**) Theoretically predicted B(*f*) plotted with the matched experimental data (zoom of **d**).

provides information on ΔG and the barrier position. CEBA (**section 6.3.2**), in contrast, allows direct determination of height of the barrier depending on the applied force, by the unfolding and folding rates to the theoretically derived barrier B(f) (**equation 6.14**). First, ln(k_0) is extracted by matching the natural logarithm of $k_{NU}(f)$ to the B(f) (**equations 6.20**, **Figure 6.4c**). In the second step the, ΔG^0 is extracted by matching the natural logarithm of $k_{UN}(f)$ to the B(f) with the just obtained value for ln(k_0) and the elastic contributions of the hairpin (**equations 6.21**, **Figure 6.4c**). We find that ΔG^0 = 49.5 ± 1.5 k_BT and ln(k_0) = 9.7 ± 0.6 corresponds to k_0 = 1.6 ± 1.3 × 10^4 s^{-1}. The value for ΔG^0 is very close to the formation energy expected from the NN model (51 k_BT)[125] and both ΔG^0 and k_0 are in good agreement with values previously obtained by optical tweezers[127]. The theoretical B(f) and the experimentally obtained data points in **Figure 6.4d** show a reasonably good overlap, which could be improved by measuring the elastic properties of the hairpin.

6

6.5 Discussion

In this chapter, we have pushed the capabilities of AFS and shown that it can be used as an efficient single-molecule force spectroscopy tool to unravel the free energy landscape of a DNA hairpin. We have performed constant force measurements, as well as dynamic force spectroscopy. We first applied the widely used Bell-Evans theory[113,120] to extract free energies from the folding and unfolding rates obtained with AFS. We obtain significantly different values for ΔG and the distances to the transition state than earlier studies of the same hairpin using optical tweezers[116,125]. One of the reasons for this could be that Bell-Evans theory assumes application of a force clamp (as in AFS) and not a distance clamp (as in optical tweezers). Bell-Evans theory states in fact that, when a force is applied along the reaction coordinate, this can be seen as adding up a linear potential (−fx) to the free energy landscape (see 6.3.1). The potential field for a distance clamp looks very different from the one for a force clamp, possibly explaining the difference between our obtained data and the data measured with optical tweezers[125]. In addition, we applied the more recently developed CEBA, which does not take into account how the load on the construct shapes the free energy landscape. Using CEBA we obtain values for ΔG^0 and k_0 that are very similar to those predicted theoretically and measured before using optical tweezers[125], confirming the insensitivity of this approach to the potential.

In principle, the Bell-Evans barrier height B_{BE} can be resolved using the temperature dependence of the unfolding rate. The unfolding rate at zero load

(k_m) can be determined using **equation 6.8** and by measuring the dependence of k_m on temperature, B_{BE} and the k_0 can be extracted. To this end, the natural logarithm of k_m should be fitted with a straight line, where the slope of the fit is equal to the ln($k0$) and the intersect equal to $-B_{BE}/k_BT$ (**equation 6.5**). Unfortunately, the current data at higher temperatures was not sufficient to accurately determine k_m; more measurements will be required to test this approach.

We show that AFS can be used to measure the dynamics of conformational changes of short constructs like the CD4 DNA hairpin studied here (60 bp ≈ 20 nm). The advantage of a short construct with very short handles is that the system is more stiff, resulting in an increased localization accuracy of the tethered microsphere. Here we measured a localization accuracy of ~3.4 nm standard deviation (see **Figure 6.2a**). Theory (**equation 2.5**) predicts that the stiffness of CD4 is ~100 higher than that of pKYB1 (~8000 bp; **section 2.4.8**), resulting in a decreased thermal noise. There is, however, another contribution to the standard deviation of the position due to thermal fluctuations, the viscous friction coefficient, which is higher close to the surface. In our current experiments, the friction coefficient 7.8×10^{-6} N s m^{-1}, for a 4.5 μm (diameter) microsphere 20 nm above the surface, while at 3 μm above the surface (the typical distance of the DNA tether plus microsphere) it is 0.5×10^{-6} N s m^{-1}[133]. The total expected standard deviation in the microsphere position due to thermal fluctuations is ~0.07 nm for the native state of the short hairpin and ~0.5 nm for the unfolded state (calculated from the stiffness of the double stranded DNA (k_{ssDNA} = 4.7×10^{-2} pN nm^{-1}, **equation 6.16**) and single stranded DNA (k_{ssDNA} = 2.9×10^{-2} pN nm^{-1}, **equation 6.16**) at 15 pN). This theoretically expected localization accuracy is considerably higher than the measured accuracy, indicating that it is limited not by thermal motion but by the tracking accuracy of our current AFS system. Localization accuracies of <1 nm have been achieved before with kHz acquisition rates[131], using higher-intensity illumination, high NA objectives and larger magnification. Consequently, in principle our instrument could be improved in the same way, which would allow measurement of faster transitions for constructs with high enough stiffness, like CD4.

One of the main strengths of AFS is the wide range of force loading rates that can be applied, here between 0.8 and 250 pN/s (see **Figure 6.4a**). While this loading rate range spans ~4 orders of magnitude, it is straightforward to further decrease the loading rate. In our experiment that was not necessary since the hairpin folding and unfolding is in equilibrium at 1 pN/s (**Figure 6.4a**)

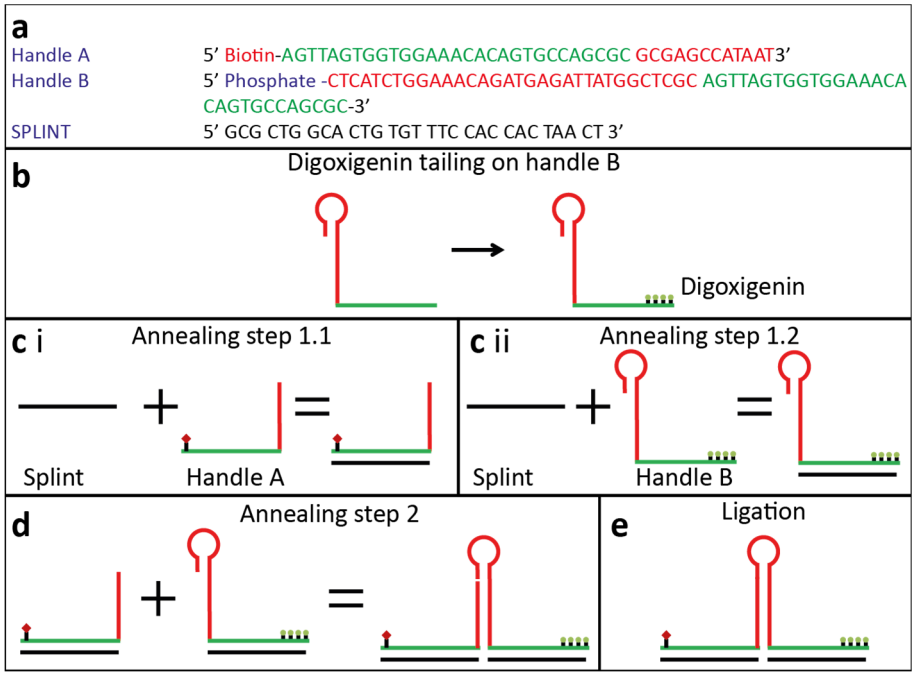
and more insight could be obtained from constant force measurements. One could, however, imagine other hairpins or molecular constructs that would require lower loading rates to reach equilibrium. In general, application of very low loading rates results in long measurement times. AFS has a clear advantage here over AFM or OT, because of its multiplexing capability. Loading rates higher than 250 pN/s could not be achieved, limited by the frame rate of the particular camera used. By using another camera, with ~100 times higher frame rate[131], higher loading rates should be possible. The limit then becomes the response time of the system, which is already visible in our data at 100 and 250 pN/s loading rate in the low force regime (orange arrows **Figure 6.4a**). Two additional factors contribute to the response time of the system: the time it takes until the acoustic standing wave reaches its final amplitude (~200 ns, **section 2.4.5**) and the response time of the tether, which is ~0.2 ms for the handles and 1 ms for the hairpin in open configuration at 15 pN (**Supplementary figure 6.6**). The response time of the tether is therefore the main contribution and it will contribute even more for lower forces. This is the reason why the forward and the backward curve do not match with a loading rate of 250 pN/s for the lower forces range (**Figure 6.4a**). This effect could be taken into account by calculating the effective force on the microsphere, or could be reduced by using smaller microspheres (**section 9.1**), or by measuring further away from the surface using very stiff handles such as for example DNA bundles based on DNA origami[132,133].

Here we have shown that AFS can be used to measure the conformation of short-handle DNA hairpins with high time and spatial resolution. To further expand the measuring capabilities of AFS, more accurate microspheres tracking systems and a camera with faster acquisition rates can be implemented. This would pave the way for single-molecule force-spectroscopy measurements on more complex systems, such as RNA and proteins, with as ultimate goal to solve the biomolecule folding problem.

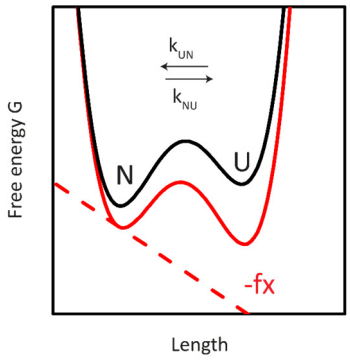
Acknowledgements

We thank Felix Ritort for sharing the CD4 hairpin construct and all the insightful scientific discussions, Anna Alemany for helping with the data analysis and Ainara Morera for teaching how to fabricate the CD4 sample.

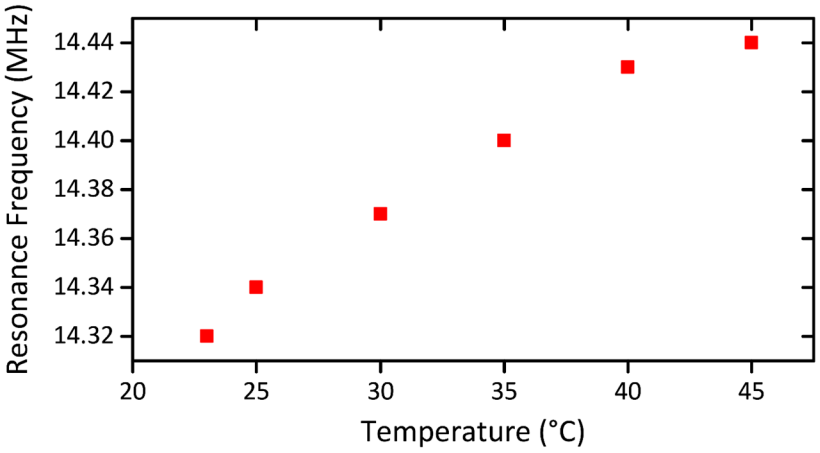
6.6 Supplementary figures



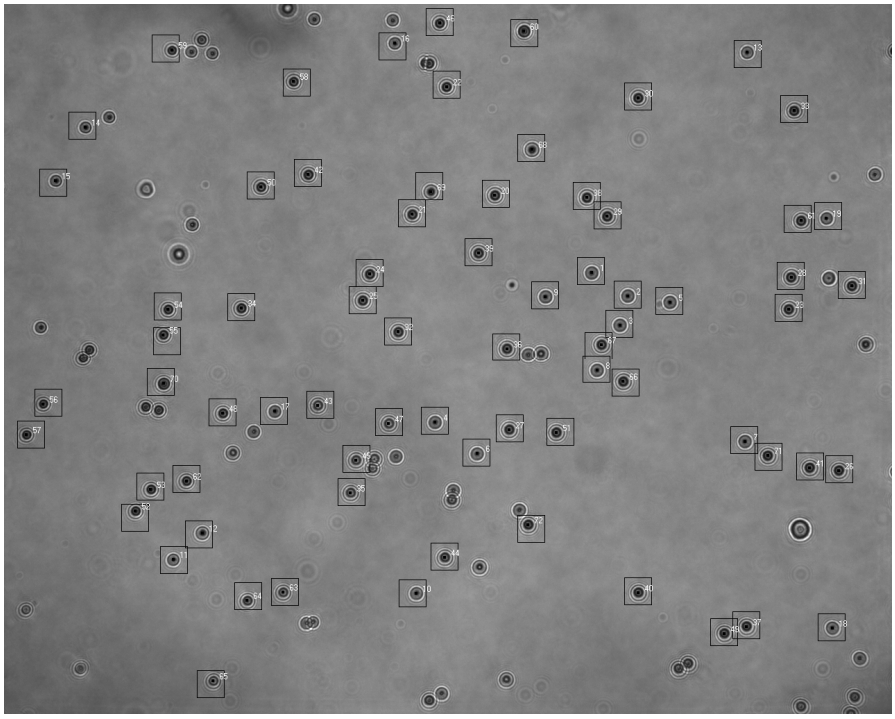
Supplementary figure 6.1 | Synthesis of the DNA hairpin
(**a**) Sequence of the 3 DNA oligos used for synthesis of the CD4 hairpin. The hairpin, the handle and the splint are colored red, green and black, respectively and color code is maintained in the rest of the figure. First handle B is labeled with digoxigenin (**b**). Next, both handles A and B are annealed to the splint separately (**c1** and **c1**), before annealing (**d**). In the last step, the backbone of the DNA hairpin is ligated (**e**).



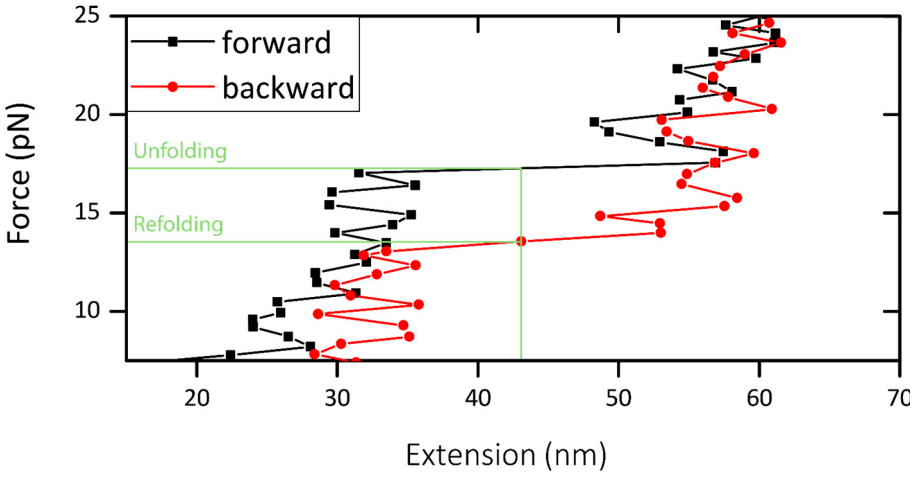
Supplementary figure 6.2 | Tilting the free-energy landscape according to Bell-Evans theory
Under the action of a mechanical force, the folding free-energy landscape of the molecule is changed by adding $-fx$ (red dotted line). This results in a tilted energy landscape (red curve) compared to the original (black curve).



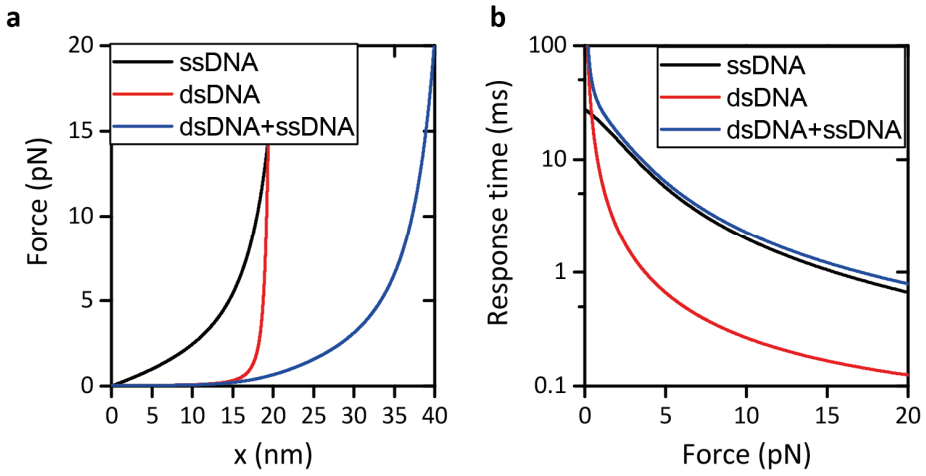
Supplementary figure 6.3 | Dependence of measured resonance frequency on temperature
The resonance frequency is measured as explained in **section 4.3.1** and plotted against the temperature, showing an increase with increasing temperature.



Supplementary figure 6.4 | A digital camera image of the field of view
The smaller diffraction patterns correspond to the surface-attached digoxigenin-coupled 3.05 µm (ø) microspheres (here number 1-20), while the bigger diffraction patterns correspond to the larger, CD4 hairpin-tethered 4.5 µm (ø) microspheres (here number 21-72).



Supplementary figure 6.5 | Example force-extension curves of a CD4 unfolding and folding event
Here, the force is ramped up with 35 pN/s (black curve and squares) and relaxed with the same loading rate (red line and dots). The first unfolding and refolding events are indicated with a green line. Data is obtained with 60 Hz acquisition rate.



Supplementary figure 6.6 | Calculating the theoretical response time
(**a**) Force-extension curve of double stranded DNA (60 bp), single stranded DNA (44 bp) and combined, related to the folded, the hairpin and the unfolded state, respectively. (**b**) Response time of the system for different forces applied (see also **section 2.4.5**), calculated by dividing the friction coefficient ($\gamma_{fax} \approx 5 \times 10^{-6}$ N s m^{-1}, for a 4.5 µm microsphere 20 nm above the surface) by the stiffness k_z in the z direction (the derivative of the force-extension curve).